DNA and Chromosomes

A genome is the full complement of genes in an organism. Genome sequencing was once a huge effort. The human genome sequencing started as early as 1980 and a draft was completed in 2001. It involved thousands of scientists. Now sequencing is relatively easy, a whole bacterial genome can be done in a weekend. The new challenge now is to relate genomes to biology.

How do you fit a genome's worth of DNA inside a cell, since for example the human genome is composed of 2900Mbp? For bacteria it isn't such a big problem. The DNA is condensed about 1000 fold. In prokaryotes, the combinations of DNA with proteins form the nucleoid, which is different than the nucleus found in eukaryotes.

1) Positively charged polyamines.
2) Numerous small proteins like H-NS, HU, IHF, FIS
3) Supercoiling of DNA by topoisomerases.

Now to get eukaryotic genome packed into cells is a much bigger problem. There are 3 billion base pairs per cell, which is equivalent to 2 meters of DNA per cell. Humans have 23 pairs of chromosomes. Each chromosome contains a single, long linear DNA molecule and associated proteins. It is important to realize that chromosomes are not always present; they form only when cells are dividing. At other times, our DNA is found in a less organized form. This association of DNA and proteins is called Chromatin. This chromatin is tightly packaged, but the DNA must remain accessible for transcription, replication and repair. There are three levels of organization of chromatin: nucleosome ("beads on a string"), 30nm condensed chromatin fiber and chromosome. Nucleosomes are the basic structural units in eukaryotic DNA and when packed together form what we have already seen, the chromatin.

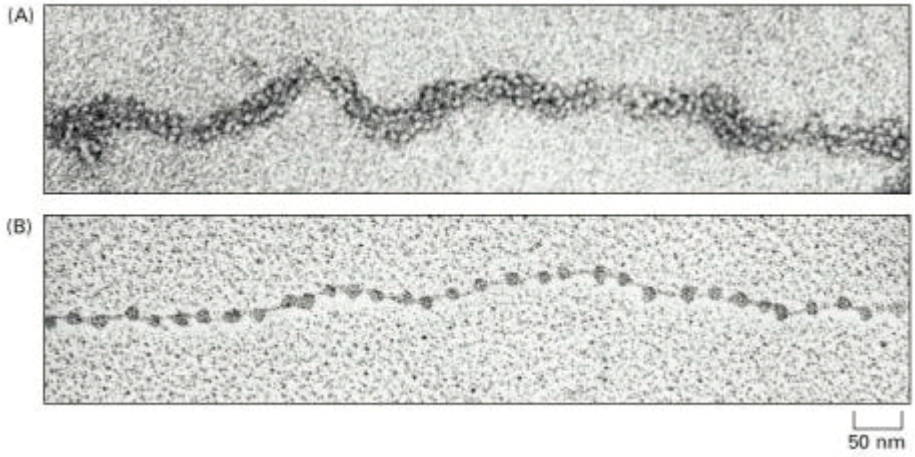


Figure 4-23. Nucleosomes as seen in the electron microscope. (A) Chromatin isolated directly from an interphase nucleus appears in the electron microscope as a thread 30 nm thick. (B) This electron micrograph shows a length of chromatin that has been experimentally unpacked, or decondensed, after isolation to show the nucleosomes.

The isolation of the nucleosome can be done experimentally. A nucleosome contains a protein core made of eight histone molecules. The nucleosome core particle can

be released from chromatin by digestion of the linker DNA with a nuclease, an enzyme that breaks down DNA. (The nuclease can degrade the exposed linker DNA but cannot attack the DNA wound tightly around the nucleosome core.) After dissociation of the isolated nucleosome into its protein core and DNA, using high concentration of salt, the length of the DNA that was wound around the core can be determined. This length of 146 nucleotide pairs is sufficient to wrap 1.65 times around the histone core.
http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=mboc4.figgrp.632

The octameric histone core is a small protein rich in lysine and arginine, which are polar/basic amino acids. The positive charge of these proteins neutralizes the negative charge of the DNA. There are four core histone proteins, H2A, H2B, H3 and H4. Therefore a pair of each of the proteins forms the octamers, in addition to one linker histone, H1.

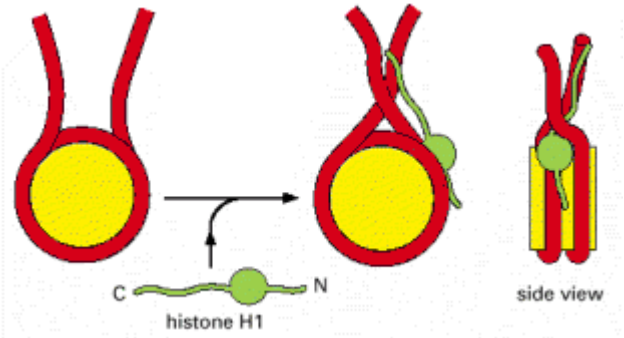


Figure 4-31. A speculative model for how histone H1 could change the path of DNA as it exits from the nucleosome. Histone H1 *(green)* consists of a globular core and two extended tails. Part of the effect of H1 on the compaction of nucleosome organization may result from charge neutralization: like the core histones, H1 is positively charged (especially its C-terminal tail), and this helps to compact the negatively charged DNA. Unlike the core histones, H1 does not seem to be essential for cell viability; in one ciliated protozoan the nucleus expands nearly twofold in the absence of H1, but the cells otherwise appear normal.

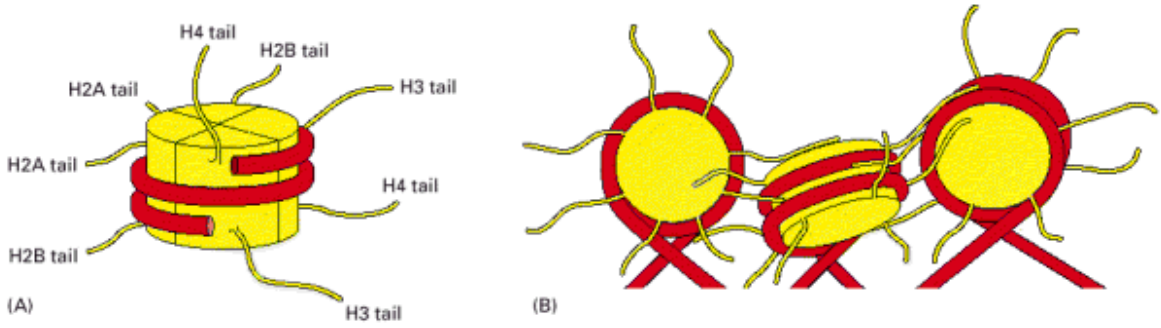


Figure 4-32. A speculative model for histone tails in the formation of the 30-nm fiber. (A) The approximate exit points of the eight histone tails, four from each histone subunit, that extend from each nucleosome. In the high-resolution structure of the nucleosome, the tails are largely unstructured, suggesting that they are highly flexible. (B) A speculative model showing how the histone tails may help to pack nucleosomes together into the 30-nm fiber. This model is based on (1) experimental evidence that histone tails aid in the formation of the 30-nm fiber, (2) the x-ray crystal structure of the nucleosome, which

showed that the tails of one nucleosome contact the histone core of an adjacent nucleosome in the crystal lattice, and (3) evidence that the histone tails interact with DNA.

Heterochromatin is a highly condensed form of chromatin. In the interphase, chromatin is attached to a chromosome scaffold. "Heterochromatic" regions of interphase chromosomes are regions where gene expression is suppressed. Heterochromatin can also be seen in centromeres, mitotic and meiotic chromosomes, and on the X chromosome in human females (Barr body).

Euchromatin in the other hand is relatively non-condensed chromatin. "Euchromatic" regions of interphase chromosomes are regions where genes may be expressed.

For a nice little animation go to http://www/dmao/org/c/index.html -> tour -> chromosome close up -> 3d animation.

Transcription

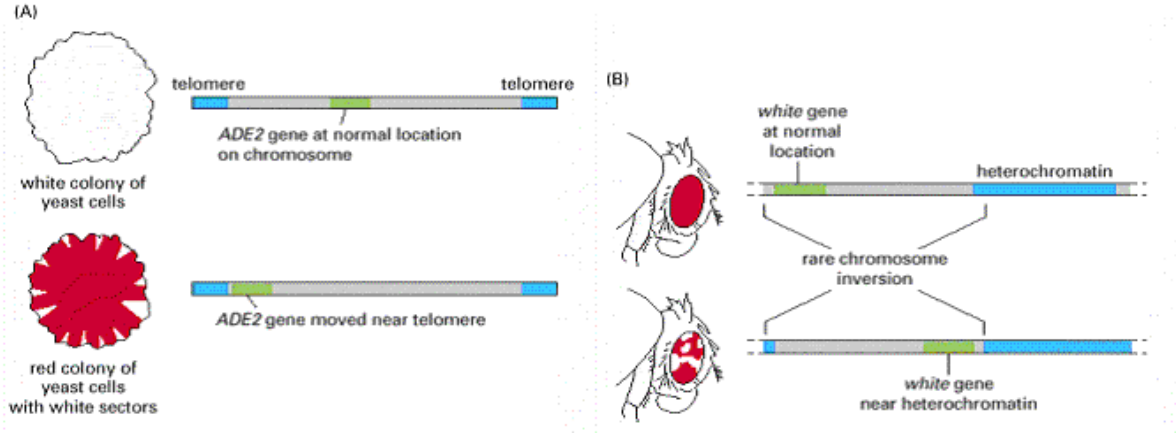
Differences in transcriptions define cells, tissues and organisms. How much you transcribe of a gene makes the differences between organisms and even cells within an organism. Now, in order to actually visualize how much difference in gene expression there actually is, between different types of cells, we can use DNA Microarray. This method provides us with a snapshot of the expression of many different genes at the same time. A clusterogram or heat map, can show how similar the genome profiles of different specimens are. There are also other ways to represent these results.

Transcriptome is the complete set of RNA transcripts produced by the genome at any one time. The Transcriptome is dynamic and changes under different circumstances (different cell, tissues, condition, time of day, etc.) due to different patterns of gene expression. The study of the Transcriptome is termed transcriptomics.

Chromatin structure affects gene expression. For example in the experiment doen with *Drasophila Melanogaster* below. Note that the *white gene* in the fruit fly, actually codes for the red color of the eyes.

Figure 4-45. Position effects on gene expression in two different eucaryotic organisms. (A) The yeast *ADE2* gene at its normal chromosomal location is expressed in all cells. When moved near the end of a yeast chromosome, which is inferred to be folded into a form of heterochromatin, the gene is no longer expressed in most cells of the population. The *ADE2* gene codes for one of the enzymes of adenine biosynthesis, and the absence of the *ADE2* gene product leads to the accumulation of a red pigment. Therefore, a colony of cells that expresses *ADE2* is *white,* and one composed of cells where the *ADE2* gene is not expressed is *red.* The *white* sectors that fan out from the middle of the *red* colony grown on an agar surface represent the descendants of cells in which the *ADE2* gene has spontaneously become active. These white sectors are thought to result from a heritable change to a less tightly packed state of chromatin near the *ADE2* gene in these cells. Although yeast chromosomes are too small to be seen under the light microscope, the chromatin structure at the ends of yeast chromosomes is thought to have many of the same structural features as the heterochromatin in the chromosomes of larger organisms. (B) Position effects can also be observed for the *white* gene in the fruit fly *Drosophila.* The *white* gene controls eye pigment production and is named after the mutation that first

identified it. Wild-type flies with a normal *white* gene *(white$^-$)* have normal pigment production, which gives them red eyes, but if the *white* gene is mutated and inactivated, the mutant flies *(white$^-$)* make no pigment and have *white* eyes. In flies in which a normal *white$^+$* gene has been moved near a region of heterochromatin, the eyes are mottled, with both *red* and *white* patches. The *white* patches represent cells in which the *white$^+$* gene has been silenced by the effects of the heterochromatin. In contrast, the red patches represent cells that express the *white$^+$* gene because the heterochromatin did not spread across this gene at the time, early in development, when the heterochromatin first formed. As in the yeast, the presence of large patches of *red* and *white* cells indicates that the state of transcriptional activity of the gene is inherited, once determined by its chromatin packaging in the early embryo.



In order to alter the access to DNA, chromatin is sometimes remodeled. Areas where the chromatin is remodeled can form loops where high-level expression of genes is possible, since that area is available for transcription.

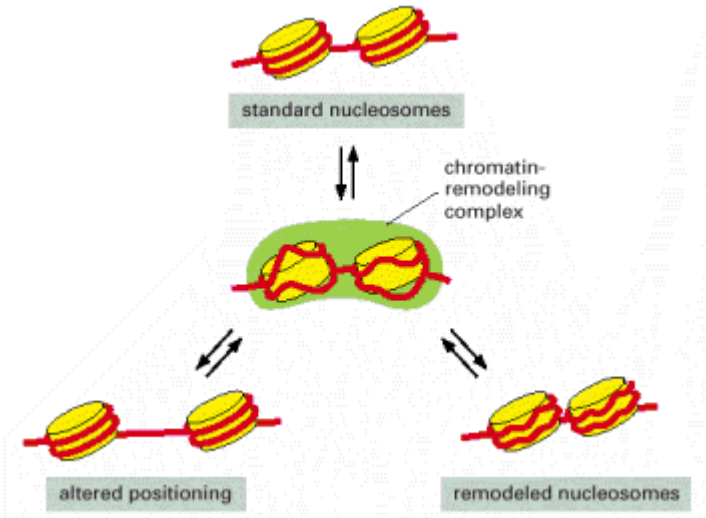


Figure 4-33. Model for the mechanism of some chromatin remodeling complexes. In the absence of remodeling complexes, the interconversion between the three nucleosomal states shown is very slow because of a high activation energy barrier. Using ATP hydrolysis, chromatin-remodeling complexes *(green)* create an activated intermediate (shown in the center of the figure) in which the histone-DNA contacts have been partly disrupted. This activated state can then decay to any one of the three nucleosomal

configurations shown. In this way, the remodeling complexes greatly increase the rate of interconversion between different nucleosomal states. The remodeled state, in which the histone-DNA contacts have been loosened, has a higher free energy level than that of standard nucleosomes and will slowly revert to the standard nucleosome conformation, even in the absence of a remodeling complex. Cells have many different chromatin remodeling complexes, and they differ in their detailed biochemical properties; for example, not all can change the position of a nucleosome, but all use the energy of ATP hydrolysis to alter nucleosome structure.

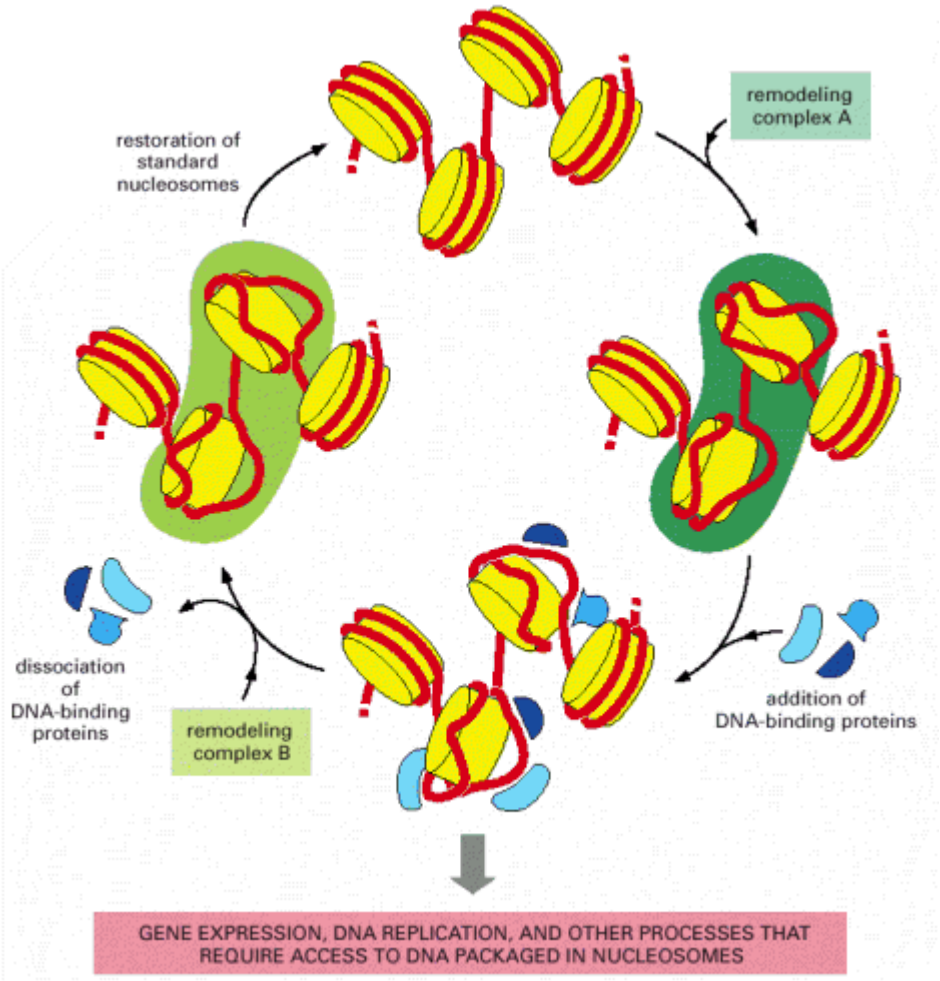


Figure 4-34. A cyclic mechanism for nucleosome disruption and re-formation. According to this model, different chromatin remodeling complexes disrupt and re-form nucleosomes, although, in principle, the same complex might catalyze both reactions. The DNA-binding proteins could function in gene expression, DNA replication, or DNA repair, and in some cases their binding could lead to the dissociation of the histone core to form nucleosome-free regions of DNA like those illustrated in Figure 4-30.

Histone tails can be modified; the N-terminal ends of a histone protrude from the core, which subjects it to the formation of covalent bonds forming a modified molecule. Shown below. I think the important thing to know here are the different molecules that can bind to the histone: Phosphate, Acetyl group, Methyl group and Ubiquitin.

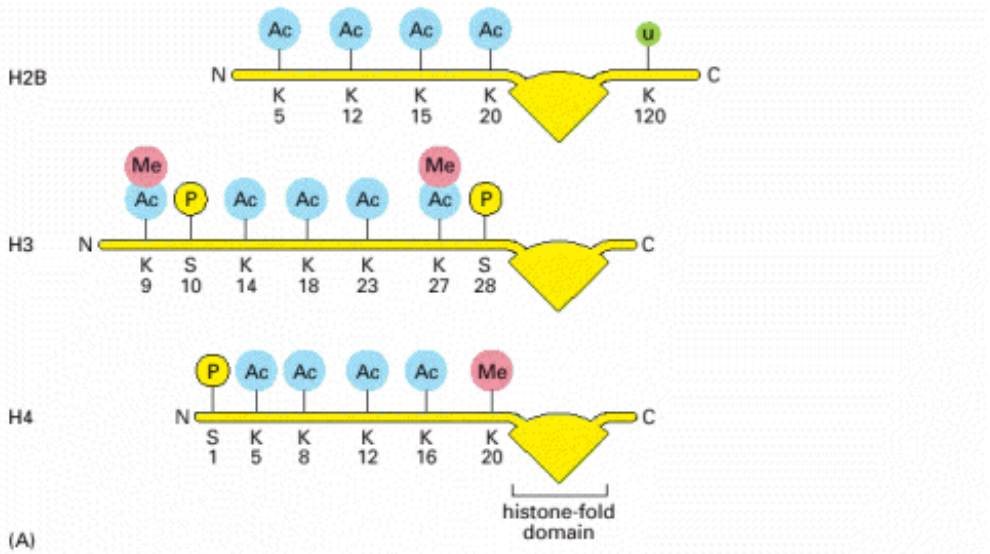
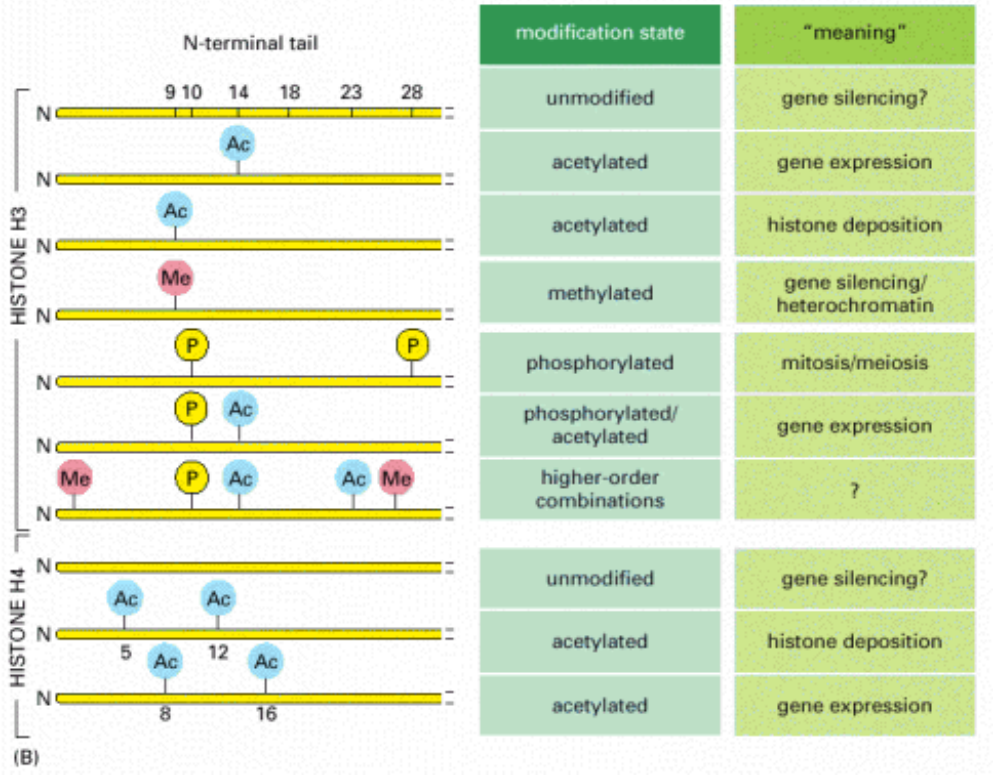


Figure 4-35. Covalent modification of core histone tails. (A) Known modifications of the four histone core proteins are indicated: Me = methyl group, Ac = acetyl group, P = phosphate, u = ubiquitin. Note that some positions (e.g., lysine 9 of H3) can be modified in more than one way. Most of these modifications add a relatively small molecule onto the histone tails; the exception is ubiquitin, a 76 amino acid protein also used in other cellular processes (see Figure 6-87). The function of ubiquitin in chromatin is not well understood: histone H2B can be modified by a single ubiquitin molecule; H2A can be modified by the addition of several ubiquitins.

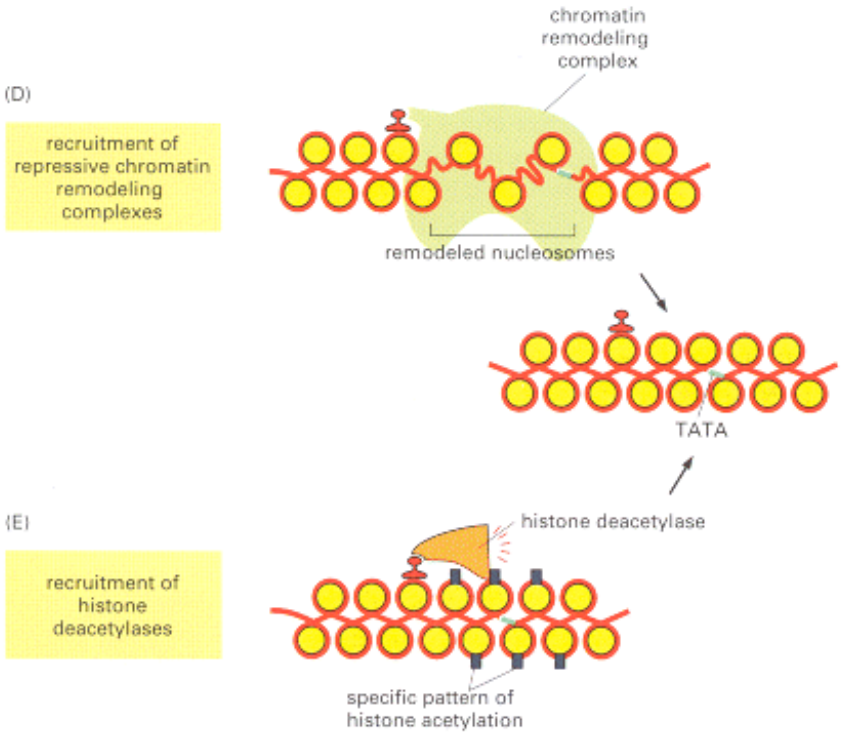


(B) A histone code hypothesis. Histone tails can be marked by different combinations of modifications. According to this hypothesis, each marking conveys a specific meaning to the stretch of chromatin on which it occurs. Only a few of the meanings of the modifications are known. In Chapter 7, we discuss the way a doubly-acetylated H4 tail is "read" by a protein required for gene expression. In another well-studied case, an H3 tail methylated at lysine 9 is recognized by a set of proteins that create an especially compact form of chromatin, which silences gene expression. The acetylation of lysine 14 of histone H3 and lysines 8 and 16 of histone H4—usually associated with gene expression—is performed by the type A histone acetylases (HATs) in the nucleus. In contrast, the acetylation of lysines 5 and 12 of histone H4 and a lysine of histone H3 takes place in the cytosol, after the histones have been synthesized but before they have been incorporated into nucleosomes; these modifications are catalyzed by type B HATs. These modified histones are deposited onto DNA after DNA replication (see Figure 5-41), and their acetyl groups are taken off shortly afterwards by histone deacetylases (HDACs). Thus, the acetylation at these positions signals newly replicated chromatin. Modification of a particular position in a histone tail can take on different meanings depending on other features of the local chromatin structure. For example, the phosphorylation of position 10 of histone H3 is associated not only with the condensation of chromosomes that takes place in mitosis and meiosis but also with the expression of certain genes. Some histone tail modifications are interdependent. For example methylation of H3 position 9 blocks the phosphorylation of H3 position 10, and vice versa.

Basically what this image above tells us is that the different modification states determines what happens to the DNA at that particular site, for example being expressed or silenced. In the chromosome, it is also possible to find sequence-specific DNA binding proteins. Among these we find sequence-specific transcription factors. Examples of one of these factors are the sequence-specific transcriptional activators, which can activate transcription, recruit histone acetylase, decondense chromatin, and facilitate the recruitment of general transcription machinery (eg. TFIID). This may occur in the following order:

1. Gene activator protein binds to chromatin → 2. Chromatin remodeling takes place → → 3. Covalent histone modification occurs → 4. Additional activator proteins bound to gene regulatory region → 5. Assembly of pre-initiation complex at the promoter → → 6. Transcription begins.

There are also sequence-specific transcriptional repressors, which repress transciption, condense chromatin and recruit histone deacetylase. These effects can be observed on the figure below.



(D) The repressor recruits a chromatin remodeling complex which returns the nucleosomal state of the promoter region to its pre-transcriptional form. Certain types of remodeling complexes appear dedicated to restoring the repressed nucleosomal state of a promoter, whereas others (for example, those recruited by activator proteins) render DNA packaged in nucleosomes more accessible (see Figure 4-34). However the same remodeling complex could in principle be used either to activate or repress transcription: depending on the concentration of other proteins in the nucleus, either the remodeled state or the repressed state could be stabilized. According to this view, the remodeling complex simply allows chromatin structure to change. (E) The repressor attracts a histone deacetylase to the promoter. Local histone deacetylation reduces the affinity of TFIID for the promoter (see Figure 7-46) and decreases the accessibility of DNA in the affected chromatin. A sixth mechanism of negative control—inactivation of a transcriptional activator by heterodimerization—was illustrated in Figure 7-26.

The implications of this control of gene expression can be seen in a practical example of the β-globin gene family. Fetal hemoglobin (contains YA and YG globin) has a higher affinity for Oxigen than the maternal hemoglobin. Therefore genes are expressed differently throughout life.